

EXHIBIT 2

1 UNITED STATES DISTRICT COURT
2 WESTERN DISTRICT OF WASHINGTON
3

4 Kaeli Garner, et al.,) Case No.
5 Plaintiffs,) 2:21-cv-00750-RSL
6 vs.)
7 Amazon.com, Inc. and Amazon.com)
8 Services LLC,)
9 Defendants.)
10

11 HIGHLY CONFIDENTIAL - ATTORNEYS' EYES ONLY

12 DEPOSITION OF
13 SERGE EGELMAN, Ph.D

14 Volume I

15 Friday, September 6, 2024

16 9:02 a.m.
17
18
19
20
21
22

23 Reported by: Michael C. Rowell, California CSR #13494

24 NCRA Registered Diplomate Reporter

25 NCRA Certified Realtime Captioner

1	BY MR. WAKEFIELD:	10:43:44
2	Q. And let me know when you have it.	10:43:44
3	A. Yep, I have it open.	10:43:45
4	Q. Great. So, I believe we're on Exhibit 4.	10:43:47
5	Do you recognize Exhibit 4 as the Privacy	10:43:53
6	Attitudes of Smart Speakers paper?	10:44:03
7	A. Yes, I'll -- I'll assume this is the true and	10:44:04
8	correct version of this paper.	10:44:05
9	Q. And I wasn't sure whether to call it a study,	10:44:07
10	a paper, or a survey. It's a -- can we agree to call it	10:44:11
11	a paper?	10:44:13
12	A. Yeah, that's fine. A publication, a paper, an	10:44:14
13	article, yeah.	10:44:17
14	Q. Okay. And does the 2019 Privacy Attitudes of	10:44:17
15	Smart Speaker Users paper accurately describe the	10:44:29
16	methodology that was used in the underlying survey?	10:44:32
17	A. Yes, I believe so, and an anonymous committee	10:44:36
18	of my peers also seems to agree with that assessment.	10:44:44
19	Q. And does -- sorry.	10:44:48
20	And does Exhibit 4, the Privacy Attitudes of	10:44:57
21	Smart Speaker Users, accurately describe the questions	10:45:01
22	that were asked in the study, in the survey?	10:45:03
23	A. Yes. I believe the survey is probably	10:45:09
24	included as an appendix.	10:45:12
25	Q. Yeah. So, and I could direct you at the	10:45:14

1	page --	10:45:17
2	A. Yeah, I have it. It starts on page 268.	10:45:18
3	Q. 268, Appendix C. It starts there, and that's	10:45:21
4	the survey instrument. Is that the -- those are the --	10:45:24
5	starting at 268, those are the questions that were	10:45:28
6	actually provided to survey participants, correct?	10:45:30
7	A. Correct.	10:45:33
8	Q. And do you agree with the conclusions and	10:45:33
9	recommendations from the Privacy Attitudes of Smart	10:45:40
10	Speaker Users paper?	10:45:45
11	A. Are you asking me to reread all of the	10:45:47
12	conclusions and recommendations?	10:45:51
13	Q. No. It's -- I mean, it's something you relied	10:45:52
14	on in your -- in your report, and it's a paper you	10:45:54
15	co-authored.	10:46:02
16	A. I mean, the -- the statements that I relied	10:46:03
17	on, I, you know, certainly -- yes, in terms of the full	10:46:06
18	set of conclusions and the recommendations that are in	10:46:11
19	this paper, which, you know, go beyond the purpose for	10:46:13
20	citing it in the report.	10:46:17
21	Whether those all still hold, again, I would	10:46:19
22	need to look at the full set of conclusions and	10:46:22
23	recommendations. But to the extent that I relied on it	10:46:23
24	in drafting my report, I -- I would, you know, agree	10:46:26
25	with those conclusions still.	10:46:28

1 Q. Do you stand by the methodology that was used 10:46:31
2 in the underlying survey? 10:46:33
3 A. Yeah, absolutely. 10:46:36
4 Q. Okay. And do you believe it was appropriate 10:46:37
5 to rely on this paper and the underlying survey in 10:46:41
6 forming expert opinions for this case? 10:46:45
7 MS. IZZO: Objection to the form. 10:46:49
8 THE WITNESS: This is a peer-reviewed paper 10:46:50
9 appearing at a, you know, reputable international 10:46:52
10 conference that, you know, it was peer-reviewed, and, 10:46:55
11 you know, yeah, absolutely. 10:47:00
12 BY MR. WAKEFIELD: 10:47:03
13 Q. This paper, Exhibit 4, was published in 2019, 10:47:05
14 correct? 10:47:12
15 A. That's correct. 10:47:13
16 Q. The underlying survey, it appears, was 10:47:13
17 conducted in February 2019; does that sound right? 10:47:18
18 A. Without scrolling back to see which page that 10:47:23
19 appears on, that -- that seems right. Usually, the 10:47:26
20 publication, you know, timeline is like six months or 10:47:29
21 so, so that -- that would seem right. 10:47:33
22 Q. And you could look specifically to page -- the 10:47:39
23 page with 254 at the top, under Summary Data. 10:47:44
24 A. Yep, that's what it says. 10:47:51
25 Q. All right. So, it says, "We conducted our 10:47:53

1 study during February 2019"; is that right? 10:47:57

2 A. That's what it says. 10:48:01

3 Q. It also says, "We piloted our study with 13 10:48:02

4 subjects and then ran the main study with 103 10:48:07

5 participants for a total of 116 respondents to our 10:48:11

6 survey"; is that correct? 10:48:15

7 A. That's correct. 10:48:18

8 Q. Based on the pilot, it looks like there were 10:48:18

9 changes made to the extension and some additional 10:48:27

10 questions; is that right? 10:48:30

11 MS. IZZO: Objection to the form. 10:48:34

12 THE WITNESS: Let me read it. 10:48:36

13 BY MR. WAKEFIELD: 10:48:40

14 Q. It's the next paragraph there under -- 10:48:41

15 A. Yeah, yeah, yeah. 10:48:41

16 Q. -- Section 5, Summary Data. 10:48:42

17 A. Yeah, that's what it says. 10:48:44

18 Q. But it also says that "...the main study and 10:48:47

19 the pilot were substantially similar, so we report the 10:48:55

20 results as one combined data set," correct? 10:48:59

21 A. Yeah, that's a common practice. 10:49:01

22 Q. Okay. And if you look at the next column on 10:49:05

23 that same page, 254, under Device Distribution. 10:49:08

24 A. Okay. 10:49:17

25 Q. It says, "Approximately two-thirds of our 10:49:17

1 participants, 69 percent, owned a smart speaker with 10:49:20
2 Amazon Alexa (such as the Echo, Echo Dot, etc.) while 10:49:25
3 the remaining 31 percent owned a Google Home or one of 10:49:30
4 its variants." 10:49:34

5 Did I read that correctly? 10:49:35

6 A. You did. 10:49:37

7 Q. Okay. And that's correct, right? 10:49:37

8 A. That's correct. 10:49:40

9 Q. All right. And it goes on in that paragraph 10:49:41
10 to say that there were no significant differences 10:49:53
11 between owners of Alexa and Google devices in their 10:49:56
12 gender or the number of interactions they had with smart 10:50:01
13 speakers, correct? 10:50:04

14 A. That's what it says. 10:50:06

15 Q. And there were also no statistical differences 10:50:07
16 between the two populations on other questions tested; 10:50:09
17 is that right? 10:50:15

18 A. That's correct. 10:50:16

19 Q. All right. So, in -- in reporting out the 10:50:16
20 results from the questions that were asked, there were 10:50:20
21 no meaningful differences in understandings or attitudes 10:50:22
22 between Alexa users and Google Home users? 10:50:27

23 A. I mean, there might have been some difference. 10:50:38
24 I mean, again, there's -- it says "see Section 6.4.1." 10:50:40
25 I would have to reread that section. 10:50:45

1 In the aggregate, yes, that's what it says. 10:50:47
2 The point is that, you know, the -- yeah, there just -- 10:50:49
3 there weren't statistically significant differences. 10:50:58
4 There might have been some nuance that we -- we talk 10:51:00
5 about in that section, but yeah, I would have to reread 10:51:03
6 that. 10:51:06
7 Q. Okay. And was this conducted by the 10:51:06
8 laboratory at Berkeley that you work with? 10:51:19
9 A. I mean, this was conducted by my laboratory. 10:51:21
10 This is, yeah, conducted by my Ph.D. student and -- and 10:51:24
11 a few undergrads who were supervised by my Ph.D. 10:51:26
12 student. 10:51:30
13 Q. Okay. And I had a question about, there -- 10:51:30
14 there are some references to "my laboratory" in your 10:51:32
15 report. So, do you -- do you run a laboratory at 10:51:36
16 Berkeley? 10:51:38
17 A. I do. 10:51:38
18 Q. Okay. And that's the laboratory that 10:51:39
19 conducted the Privacy Attitudes of Smart Speaker Users 10:51:41
20 study? 10:51:45
21 A. That's correct. 10:51:46
22 Q. Okay. Were you personally involve -- involved 10:51:47
23 in the design of the underlying survey? 10:51:51
24 A. I was. 10:51:54
25 Q. Okay. And did you review the survey questions 10:51:55

1	before they were used?	10:51:59
2	A. Absolutely.	10:52:00
3	Q. All right. Did you personally participate in	10:52:01
4	performing the survey?	10:52:05
5	A. Define "performing."	10:52:07
6	Q. Identifying people in the -- in the universe,	10:52:10
7	administering the -- the questions, writing it.	10:52:17
8	A. I -- as -- as the senior author on the paper,	10:52:23
9	I helped with the design of the study and mentored the	10:52:26
10	students and, you know, checked their work. But in	10:52:30
11	terms of, you know, meeting with study participants or	10:52:32
12	doing the actual recruiting, that -- well, that's a task	10:52:37
13	that students generally do.	10:52:39
14	Q. And was that the case with this survey?	10:52:44
15	A. I mean, that's the case with almost every	10:52:47
16	academic publication.	10:52:50
17	Q. All right. But just to -- just to be clear	10:52:51
18	that we have the answer to the question, that was true	10:52:54
19	with this --	10:52:57
20	A. Yes.	10:52:57
21	Q. -- this study too?	10:52:59
22	Okay. And the survey was of owners of smart	10:52:59
23	speakers, correct?	10:53:14
24	A. Yes.	10:53:18
25	Q. Did you determine if they had registered the	10:53:20

1 devices?

10:53:27

2 A. I would actually need to read the -- reread
3 the paper to give you that level of specificity, but it
4 was certainly people who had access to their accounts on
5 those devices.

10:53:35

10:53:37

10:53:40

10:53:44

6 Q. And let's go to that. So, as I understand it,
7 there was a browser extension they could use that they
8 could then have a controlled access to the recordings
9 that would be available to them through the -- through
10 Amazon.com, through their account?

10:53:45

10:53:50

10:53:55

10:54:03

10:54:07

11 A. Yeah. To preserve privacy and for data
12 minimization purposes, since we didn't want to, you
13 know, collect the raw audio to, you know, play for them
14 for this study, we wrote a browser extension so that,
15 you know, they could log into their device, and it would
16 show them clips of audio that had been recorded without
17 us needing to actually possess that audio or even listen
18 to it.

10:54:09

10:54:14

10:54:17

10:54:20

10:54:24

10:54:25

10:54:29

10:54:32

19 Q. How were the specific recordings selected?

10:54:34

20 A. I need to go -- I think I would need to go
21 look at the paper again to answer that.

10:54:39

10:54:42

22 Q. Okay. Do you have a recollection of whether
23 it was random?

10:54:44

10:54:45

24 A. I don't remember honestly.

10:54:47

25 MS. IZZO: Objection to the form.

10:54:49

1	BY MR. WAKEFIELD:	10:54:51
2	Q. Okay. So, in order to -- to use the browser	10:54:52
3	extension and participate in the survey, a survey	10:55:01
4	participant had to have access to the account, correct?	10:55:06
5	A. I believe that was the case, yes.	10:55:10
6	Q. Okay. Are you familiar with the concept of a	10:55:12
7	control in a survey?	10:55:28
8	A. It depends on what you're talking about, but	10:55:31
9	yes.	10:55:33
10	Q. Yeah. So, I mean, obviously, there's --	10:55:34
11	"control" can mean different things, but in the context	10:55:37
12	of survey methodology, what is your understanding of a	10:55:39
13	control?	10:55:42
14	A. I mean, there are several different	10:55:43
15	definitions. I guess it depends on what you're talking	10:55:45
16	about specifically.	10:55:48
17	Q. So, why don't you walk me through the -- the	10:55:49
18	different kinds of controls. I mean, I'm familiar with	10:55:53
19	you could have a control group. You could have control	10:55:55
20	questions. So --	10:55:59
21	A. Sure. So I --	10:56:02
22	Q. -- why don't you -- why don't you walk me	10:56:03
23	through the controls that can be used in a survey.	10:56:04
24	A. I mean, again, you're the one who's using the	10:56:08
25	term, so given the multiple definitions, I -- I would	10:56:11

1 like to know what definition you're using if you want me 10:56:13
2 to explain it. Yeah. 10:56:16

3 Q. Okay. Well, so, was there any group of 10:56:18
4 differently situated users that you tested to compare 10:56:33
5 against the -- the intended survey participants? 10:56:39

6 A. One, I would probably need to reread the paper 10:56:46
7 to answer that definitively, but based on what you've 10:56:48
8 just asked, from my understanding, it doesn't sound like 10:56:52
9 that would have been relevant to this study. 10:56:56

10 Q. Okay. What about control questions? 10:56:58
11 So, for example, a proposed answer that is 10:57:09
12 meant to test -- in a multiple-choice survey -- 10:57:16

13 A. Sure. 10:57:16

14 Q. -- a proposed answer that's meant to weed out 10:57:20
15 people who don't have any understanding or -- or to see 10:57:24
16 the extent to which people have ideas that are wrong. 10:57:26

17 A. Yeah. So, those are some -- those are often 10:57:29
18 referred to as attention check questions or 10:57:31
19 comprehension questions and can be used to filter out, 10:57:33
20 yes, people who didn't read the instructions or don't 10:57:36
21 understand them. 10:57:39

22 You could have a question where, you know, if 10:57:40
23 someone had read the -- you know, read the instructions, 10:57:42
24 the answer would be obvious, and then you can filter out 10:57:44
25 all of the people who didn't answer that question 10:57:46

1 correctly. That's one type of control. 10:57:48

2 We generally do use those in surveys, and we 10:57:50
3 also use various other methods to detect cheating in 10:57:54
4 surveys, you know, people who are just clicking random 10:57:59
5 buttons to get to the end. And, you know, there are a 10:58:03
6 lot of techniques that are fairly common. One colleague 10:58:06
7 and frequent co-author has written a lot on this 10:58:10
8 subject. So, I think I have a pretty good understanding 10:58:13
9 of it. 10:58:16

10 Q. Okay. And do you know if such an approach was 10:58:17
11 used in this case? There's a comprehend -- an attention 10:58:24
12 check question or comprehension question? 10:58:32

13 A. I would need to reread all of the survey 10:58:35
14 questions to answer that. But, you know, generally 10:58:37
15 speaking, we -- we usually ask questions like that or 10:58:40
16 use other methods to -- to filter out -- 10:58:44

17 I mean, that said, the current research on 10:58:47
18 this has shown that, you know, just looking at -- well, 10:58:50
19 I need to -- I need to go back to see how we recruited 10:58:54
20 people. But generally, yes, we use techniques to filter 10:58:57
21 out people who did not pay attention or did not 10:59:01
22 understand the questions. 10:59:04

23 Q. Do you have access to the underlying survey 10:59:06
24 data from this survey? 10:59:17

25 A. I honestly have no idea. I mean, this was 10:59:20

1 conducted five and a half years ago. Generally, when 10:59:22
2 the paper is published, the data is of little value 10:59:25
3 anymore, and there are also, you know, university 10:59:29
4 retention policies that we need to deal with. And so, I 10:59:34
5 mean, ultimately, the raw data belongs to the Regents of 10:59:39
6 the University of California. 10:59:42

7 Where it actually resides at this moment in 10:59:47
8 time, I honestly couldn't tell you, because, as you 10:59:50
9 know, this was performed five and a half years ago. 10:59:52

10 Q. Okay. Did you do anything to determine where 10:59:55
11 the data resides, if it's still around? 10:59:58

12 A. I have not looked into that -- 11:00:01

13 MS. IZZO: Objection to the form. 11:00:03

14 THE WITNESS: I have not looked into that for 11:00:04
15 this study, no. 11:00:06

16 BY MR. WAKEFIELD: 11:00:07

17 Q. Okay. And did you ask anyone at the 11:00:07
18 University if you could access that information? 11:00:13

19 A. No. 11:00:16

20 Q. Okay. Did you make any effort to try to get 11:00:17
21 the underlying data in response to the document requests 11:00:21
22 in this case? 11:00:24

23 A. It -- 11:00:27

24 MS. IZZO: Objection to the form. 11:00:27

25 THE WITNESS: It didn't seem necessary, given 11:00:28

1 that the paper is publicly available. The paper was 11:00:30

2 peer-reviewed; it's a peer-reviewed publication. 11:00:33

3 Generally, you know, once it passes peer review and is 11:00:35

4 published, no one really pays attention to the raw 11:00:39

5 underlying data anymore. 11:00:42

6 BY MR. WAKEFIELD: 11:00:45

7 Q. The survey questionnaire had places where 11:00:45

8 survey participants could provide answers to open-ended 11:00:47

9 questions. 11:00:51

10 Do you recall that? 11:00:53

11 A. I assume that's the case. It's generally a 11:00:54

12 good, you know, good practice to have some open-ended 11:00:57

13 questions. 11:01:03

14 Q. So, for example, if you look at page 270 11:01:04

15 within the survey questionnaire, toward the bottom of 11:01:10

16 the right column, it says, "In the future, do you intend 11:01:14

17 to take any steps to protect your privacy when using 11:01:21

18 your device? What do you plan to do?" 11:01:25

19 Do you see that? 11:01:28

20 A. Yep, yes. There are a couple of open-ended 11:01:29

21 questions here. 11:01:32

22 Q. Yeah. And did the paper, Exhibit 4, provide 11:01:34

23 all of the responses to those open-ended questions? 11:01:37

24 A. The paper almost certainly wouldn't have 11:01:41

25 provided all of the responses verbatim, given that there 11:01:43

1 were over a hundred responses. The -- generally, this 11:01:46
2 type of research, you look at the breadth of responses 11:01:50
3 and describe them, but, you know, publishing the raw 11:01:54
4 data generally doesn't occur. 11:01:59

5 Q. Right. So, if we wanted to know the full set 11:02:01
6 of answers that users provided in response to this 11:02:09
7 survey, we would need to have access to the underlying 11:02:12
8 data, right? 11:02:15

9 MS. IZZO: Objection to the form. 11:02:17

10 THE WITNESS: I mean, generally, the case with 11:02:19
11 qualitative data like this, again, after the paper, 11:02:20
12 after those responses have been coded and the broad -- 11:02:27
13 you know, usually, you report on the breadth of 11:02:34
14 responses, because many of the responses are going to be 11:02:37
15 overlapping, and so there's a whole qualitative coding 11:02:40
16 process. 11:02:43

17 There -- you know, there's certainly a lot of 11:02:44
18 method -- you know, methods books that will describe 11:02:46
19 those processes, but there are certainly accepted 11:02:48
20 processes for coding that data, and so you can distill 11:02:51
21 it down to aggregate responses, which is generally what 11:02:54
22 people care about with this type of research. 11:02:58

23 You know, certainly there are -- you know, 11:03:00
24 it's common to have quotes here and there, if there's 11:03:03
25 something that's particularly useful, but by and large, 11:03:05

1 with this type of research, the interest is in looking 11:03:08
2 at what percentage said X and what percentage said Y, 11:03:12
3 without actually disclosing all of the responses. 11:03:15

4 Q. Okay. But, again, if we actually wanted to 11:03:18
5 see the verbatim responses, we'd have to look at the 11:03:21
6 underlying data, because the verbatim responses are not 11:03:24
7 in the paper, correct? 11:03:26

8 A. Well, the verbatim responses, no. I mean, if 11:03:30
9 you want the underlying data, subpoena the University of 11:03:32
10 California. 11:03:35

11 Q. Okay. But, again, you didn't even try to get 11:03:37
12 any of the data for this case, did you? 11:03:42

13 A. As I stated -- 11:03:45

14 MS. IZZO: Objection to the form. 11:03:47

15 THE WITNESS: As I already stated, I don't 11:03:48
16 know where that data is, and it's irrelevant, because 11:03:50
17 once the paper is published, there's really no need to 11:03:52
18 go back to that data. 11:03:56

19 BY MR. WAKEFIELD: 11:03:57

20 Q. So, because you determined it's irrelevant, 11:03:58
21 you decided not to try to get it; is that fair? 11:03:59

22 MS. IZZO: Objection to the form, misstates 11:04:02
23 the witness's testimony. 11:04:03

24 THE WITNESS: Yeah, no, that's not accurate. 11:04:05

25 /// 11:04:08

1 BY MR. WAKEFIELD:

11:04:08

2 Q. Did you try to get the data?

11:04:08

3 A. Oh, I just -- I wouldn't have seen any need
4 given that, you know, I wouldn't have tried to get the
5 data for any of the other papers that I didn't
6 co-author.

11:04:10

11:04:13

11:04:15

11:04:18

7 Once the -- you know, once it's passed the
8 peer-review process and is published, there's really no
9 need to look at the raw data, especially for this
10 purpose. And even in this particular study, I mean,
11 this is one study of many that are cited in the report
12 that all are -- you know, they're not in conflict with
13 each other. So, you could certainly rely on other
14 studies in here that corroborate the findings.

11:04:18

11:04:22

11:04:25

11:04:27

11:04:30

11:04:33

11:04:38

11:04:40

15 Q. So, did you try to get the underlying data for
16 the Privacy Attitudes of Smart Speaker Users, yes or no?

11:04:44

11:04:47

17 A. As I already stated, no, I haven't looked at
18 the underlying data for any of the studies that I cited
19 in this paper, because there wasn't a need.

11:04:51

11:04:55

11:04:57

20 Q. Yeah. Just to clarify, I wasn't asking if you
21 looked at it. I was asking if you tried to get it.

11:04:59

11:05:01

22 Did you try to get it for this case?

11:05:04

23 A. No. I didn't see the need.

11:05:06

24 Q. Okay, thank you.

11:05:07

25 At page 268 of the questionnaire, within

11:05:14